



2002 Economic Census: Survey of Business Owners Methodology

Appendix C. Methodology

[Sources](#)
 [Industry classifications](#)
 [Sampling and estimation](#)
 [Reliability of estimates](#)
 [Comparability 2002-1997](#)
 [Return to introductory text](#)

SOURCES OF THE DATA

The 2002 Survey of Business Owners (SBO) was conducted by mail. One of two census forms was mailed to a random sample of businesses selected from a list of all firms operating during 2002 with receipts of \$1,000 or more, except those classified in the following NAICS industries:

- crop and animal production (NAICS 111, 112)
- scheduled air transportation (NAICS 4811, part)
- rail transportation (NAICS 482)
- postal service (NAICS 491)
- funds, trusts, and other financial vehicles (NAICS 525), except real estate investment trusts (NAICS 525930)
- religious, grantmaking, civic, professional, and similar organizations (NAICS 813)
- private households (NAICS 814), and
- public administration (NAICS 92).

The lists of all firms (or universe) are compiled from a combination of business tax returns and data collected on other economic census reports. The Census Bureau obtains electronic files from the Internal Revenue Service (IRS) for all companies filing IRS Form 1040, Schedule C (individual proprietorship or self-employed person); 1065 (partnership); any one of the 1120 corporation tax forms; and 941 (Employer's Quarterly Federal Tax Return). The IRS provides certain identification, classification, and measurement data for businesses filing those forms.

For most firms with paid employees, the Census Bureau also collected employment, payroll, receipts, and kind of business for each plant, store, or physical location during the 2002 Economic Census.

The report forms used to collect information are available at www.census.gov/csd/sbo/index.html.

The SBO is conducted on a company or firm basis rather than an establishment basis. A company or firm is a business consisting of one or more domestic establishments that the reporting firm specified under its ownership or control at the end of 2002. Firms were instructed to return their completed report form by mail. Two report form remails were conducted at one-month intervals to all delinquent respondents. A telephone follow-up was conducted to obtain a subset of information from selected firms that failed to return their report form. The returned forms underwent extensive review and computer processing. All reports were geographically coded, data-keyed, and edited. The editing process identified records with significant problems and firms were contacted for correction resolution. Corrections were performed interactively using standard procedures.

The data were then tabulated by NAICS, subjected to further data analysis, and the resulting corrections applied to individual computer records. Corrected tabulations were then produced for the final published reports.

A more detailed examination of census methodology is presented in the *History of the 2002 Economic Census* at www.census.gov/econ/www/history.html.

INDUSTRY CLASSIFICATION OF FIRMS

The classification for all establishments are based on the *North American Industry Classification System, United States, 2002*, manual. The kind-of-business or industry classification codes for the SBO are obtained from the 2002 Economic Census. More information on the industry classification codes is included in the Industry Classifications and Relationship to Historical Industry Classifications sections in the [Introductory Text](#).

SAMPLING AND ESTIMATION METHODOLOGIES

Sampling. To design the 2002 SBO sample, the Census Bureau used the following sources of information to estimate the probability that a business was minority- or women-owned:

- Administrative data from the Social Security Administration.
- Lists of minority- and women-owned businesses published in syndicated magazines, located on the Internet, or disseminated by trade or special interest groups.
- Word strings in the company name indicating possible minority ownership (derived from 1997 survey responses).
- Racial distributions for various state-industry classes (derived from 1997 survey responses) and racial distributions for various ZIP Codes.
- Gender, race, and Hispanic or Latino origin responses of a single-owner business to an SBO previous survey or to the 2000 Decennial Census.

These probabilities were then used to place each firm in the SBO universe in one of nine frames for sampling:

- American Indian
- Asian
- Black or African American
- Hispanic
- Non-Hispanic white men
- Native Hawaiian and Other Pacific Islander
- Other (a different race was supplied as a write-in to another source)
- Publicly owned
- Women

The SBO universe was stratified by state, industry, frame, and whether the company had paid employees in 2002. The Census Bureau selected large companies, including those operating in more than one state, with certainty. These companies were selected based on volume of sales, payroll, or number of paid employees. All certainty cases were sure to be selected and represented only themselves (i.e., had a selection probability of one and a sampling weight of one). The certainty cutoffs varied by sampling stratum, and each stratum was sampled at varying rates, depending on the number of firms in a particular industry in a particular state. The remaining universe was subjected to stratified systematic random sampling.

A firm selected into the sample was mailed one of two questionnaires. The Census Bureau sent the SBO-1 questionnaire to partnerships and corporations. The businesses were asked to report the percentage of ownership, gender, Hispanic or Latino origin, race, and several characteristic questions (e.g., age, education level) for each of the three largest percentage owners. The SBO-2 questionnaire was used for sole proprietors and self-employed individuals. The businesses were asked essentially the same information as asked on the SBO-1, but limited to two owners.

Treatment of Nonresponse. Approximately 81 percent of the 2.3 million businesses in the SBO sample responded to the survey. Data from the 1997 survey were used for businesses in both the 1997 and 2002 samples. For the remaining nonrespondents, gender, Hispanic or Latino origin, and race were imputed from donor respondents with similar characteristics (state, industry, employment status, size, and sampling frame).

Tabulation. Business ownership is defined as having 51 percent or more of the stock or equity in the business and is categorized by:

- Gender: Male; Female; or Equally Male-/Female-Owned
- Ethnicity: Hispanic or Latino Origin; Not Hispanic or Latino Origin
- Race: White; Black or African American; American Indian or Alaska Native; Asian; Native Hawaiian or Other Pacific Islander

Firms equally male-/female-owned were counted and tabulated as a separate category.

Businesses could be tabulated in more than one racial group. This can result because:

- a. the sole owner reported more than one race;
- b. the majority owner reported more than one race;
- c. a majority combination of owners reported more than one race.

For the tabulations by gender, Hispanic or Latino origin, and race, the data for each firm in the SBO sample were weighted by the reciprocal of the firm's probability of selection.

RELIABILITY OF ESTIMATES

The figures shown in this report are, in part, estimated from a sample and will differ from the figures which would have been obtained from a complete census. Two types of possible errors are associated with estimates based on data from sample surveys: sampling errors and nonsampling errors. The accuracy of a survey result depends not only on the sampling errors and nonsampling errors measured, but also on the nonsampling errors not explicitly measured. For particular estimates, the total error may considerably exceed the measured errors. The following is a description of the sampling and nonsampling errors associated with this tabulation.

Sampling variability. The particular sample used for this survey is one of a large number of all possible samples of the same size that could have been selected using the same sample design. Estimates derived from the different samples would differ from each other. The relative standard error is a measure of the variability among the estimates from all possible samples. The estimated relative standard errors presented in the tables estimate the sampling variability, and thus measure the precision with which an estimate from the particular sample selected for this survey approximates the average result of all possible samples. Relative standard errors are applicable only to those published cells in which sample cases are tabulated. A relative standard error is an expression of the standard error as a percent of the quantity being estimated.

The sample estimate and an estimate of its relative standard error can be used to estimate the standard error and then construct interval estimates with a prescribed level of confidence that the interval includes the average results of all samples. To illustrate, if all possible samples were surveyed under essentially the same condition, and estimates calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average value of all possible samples.
2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average value of all possible samples.

Thus, for a particular sample, one can say with specified confidence that the average of all possible samples is included in the constructed

interval.

Example of a confidence interval. Suppose the estimate of women-owned employer businesses in construction is 51,707 and the estimated relative standard error is 2 percent. An approximate 90-percent confidence interval is $51,707 \pm (1.6 \times 1,034)$ or $51,707 \pm 1,654$ or 50,053 to 53,361.

Nonsampling errors. All surveys and censuses are subject to nonsampling errors. Nonsampling errors are attributable to many sources, including the inability to obtain information for all cases in the universe, imputation for missing data, data errors and biases, mistakes in recording or keying data, errors in collection or processing, and coverage problems.

Explicit measures of the effects of these nonsampling errors are not available. However, it is believed that most of the important operational and data errors were detected and corrected through an automated data edit designed to review the data for reasonableness and consistency. Quality control techniques were used to verify that operating procedures were carried out as specified.

COMPARABILITY OF THE 2002 AND 1997 SBO DATA

The following changes were made in survey methodology in 2002 which affect comparability with past reports:

1. The 1997 Surveys of Minority- and Women-Owned Business Enterprises (SMOBE/SWOBE) form that was mailed to sole proprietors or self-employed individuals who were single filers or who filed joint tax returns instructed the respondent to mark one box that best described the gender, Spanish/Hispanic/Latino origin, and race of the primary owner(s). The gender question included an equal male/female ownership option. The 2002 SBO form that was mailed to sole proprietors or self-employed individuals who were single filers or who filed a joint tax return instructed the respondent to provide the percentage of ownership for each owner and the gender of the owner(s). The equal male/female ownership option was eliminated.

The form that corporations/partnerships received in 1997 requested the percentage of ownership by gender of the owners. In 2002, a business was asked to report the percentage of ownership and gender for each of the three largest percentage owners.

Male/female ownership of a business in both 1997 and 2002 was based on the gender of the person(s) owning the majority interest in the business. However, in 2002, equally male/female ownership was based on equal shares of interest reported for businesses with male and female owners. Businesses equally male-/female-owned were tabulated and published as a separate entity in both 1997 and 2002.

The 1997 SWOBE/SMOBE forms may be viewed at www.census.gov/epcd/www/pdf/97cs/mb1.pdf (corporations/partnerships) or at www.census.gov/epcd/www/pdf/97cs/mb2.pdf (sole proprietors or self-employed individuals).

The 2002 SBO forms may be viewed at www.census.gov/csd/sbo/sbo1.pdf (corporations/partnerships) or at www.census.gov/csd/sbo/sbo2.pdf (sole proprietors or self-employed individuals).

2. The Hispanic or Latino origin and racial response categories were updated in 2002 to meet the latest Office of Management and Budget (OMB) guidelines. There were nineteen check-box response categories and four write-in areas on the 2002 SBO questionnaire, compared to the twenty check-box response categories and five write-in areas on the 1997 SMOBE/SWOBE.

The Hispanic or Latino origin of business ownership was defined as two groups:

- Hispanic or Latino
- Not Hispanic or Latino

Four Hispanic subgroups were used on the survey questionnaires: Mexican, Mexican American, Chicano; Puerto Rican; Cuban; and Other Spanish/Hispanic/Latino.

The 2002 SBO question on race included fourteen separate response categories and two areas where respondents could write in a more specific race. The response categories and write-in answers were combined to create the following five standard OMB race categories:

- American Indian and Alaska Native
- Asian
- Black or African American
- Native Hawaiian and Other Pacific Islander
- White

Response check boxes were added for "Samoa" and "Guamanian or Chamorro."

The check box for "Some Other Race" and the corresponding write-in area provided in 1997 were deleted.

If the "American Indian and Alaska Native" race category was selected, the respondent was instructed to print the name of the enrolled or principal tribe.

In 1997, sole proprietors or self-employed individuals who were single filers or who filed a joint tax return were asked to mark a box to indicate the Spanish/Hispanic/Latino origin of the primary owner(s) and to mark the one box that best described the race of the primary owner(s). In 2002, they were asked to provide the percentage of ownership for the primary owner(s), his/her Spanish/Hispanic/Latino origin, and to select one or more race categories to indicate what the owner considers himself/herself to be.

The form that corporations/partnerships received in 1997 requested the percentage of ownership by Spanish/Hispanic/Latino origin and race of the owners. In 2002, a business was asked to report the percentage of ownership, Spanish/Hispanic/Latino origin, and race for each of the three largest owners, allowing them to mark one or more races to indicate what the owner considers himself/herself to be. The 2002 SBO was the first economic census in which each owner could self-identify with more than one racial group, so it was possible for a business to be classified and tabulated in more than one racial group.

Business ownership in both 1997 and 2002 was based on the Hispanic or Latino origin and race of the person(s) owning majority interest

in the business; however, in 2002, multiple-race reporting by the owner(s) could affect where a business was classified.

Note: In the 2000 population census, 2.4 percent of the population reported more than one race.

3. The Native Hawaiian- and Other Pacific Islander-Owned Firms report is new for 2002. Previously, estimates for this group of business owners were included in the Asian- and Pacific Islander-Owned Firms report for some tables (at the U.S., state, and metropolitan area by kind of business level). However, estimates at the county, place, and size of firm (employment, receipts) levels provided only the total number of businesses classified as Asian- and Pacific Islander-owned, with no detailed estimates by subgroup. Therefore, particular care should be taken in comparing the estimates for Asian-owned firms and/or Native Hawaiian- and Pacific Islander-owned firms from 1997 to 2002.

Source: U.S. Census Bureau, 2002 Economic Census

Last revised: January 30, 2006

Census Bureau Links: [Home](#) · [Search](#) · [Subjects A-Z](#) · [FAQs](#) · [Data Tools](#) · [Catalog](#) · [Census 2000](#) · [Quality](#) · [Privacy Policy](#) · [Contact Us](#)

U S C E N S U S B U R E A U
Helping You Make Informed Decisions

Page Last Modified: January 30, 2006